Title: Teaching Through Interactions in Secondary School Classrooms: Revisiting the Factor Structure and Practical Application of the Classroom Assessment Scoring System-Secondary

Authors: Christopher A. Hafen, Bridget K. Hamre, Joseph P. Allen, Courtney A. Bell, Drew H. Gitomer, & Robert C. Pianta

Publication Date: June 2014 (Online)

Article

Journal of Early Adolescence
1–30

DOI: 10.1177/0272431614537117
jea.sagepub.com

Teaching Through Interactions in Secondary School Classrooms: Revisiting the Factor Structure and Practical Application of the Classroom Assessment Scoring System–Secondary

Christopher A. Hafen[1], Bridget K. Hamre[1], Joseph P. Allen[1], Courtney A. Bell[2], Drew H. Gitomer[3], and Robert C. Pianta[1]

Abstract

Valid measurement of how students' experiences in secondary school classrooms lead to gains in learning requires a developmental approach to conceptualizing classroom processes. This article presents a potentially useful theoretical model, the Teaching Through Interactions framework, which posits teacher-student interactions as a central driver for student learning and that teacher-student interactions can be organized into three major domains. Results from 1,482 classrooms provide evidence for distinct emotional, organizational, and instructional domains of teacher-student

[1]University of Virginia, Charlottesville, USA
[2]Educational Testing Service, Princeton, NJ, USA
[3]Rutgers University, New Brunswick, NJ, USA

Corresponding Author:
Christopher A. Hafen, Center for Advanced Study of Teaching and Learning, University of Virginia, 350 Old Ivy Way, Charlottesville, VA 22903, USA.
Email: chris.hafen@gmail.com

interaction. It also appears that a three-factor structure is a better fit to observational data than alternative one- and two-domain models of teacher-student classroom interactions, and that the three-domain structure is generalizable from 6th through 12th grade. Implications for practitioners, stakeholders, and researchers are discussed.

Keywords

academic achievement, school context, teachers/teacher-adolescent relationship, learning/mathematics/reading, middle school

Recent educational reforms place new emphasis on the important role of teachers in promoting student achievement. States and districts across the country are developing and testing new models for evaluating teachers that provide fair and valid approaches to documenting teachers' performance. Most of these newer systems include multiple metrics of teacher effectiveness, including value-added scores, student reports, and observation (Kane et al., 2012). Observational methods of assessing teacher performance have particular relevance as they can provide actionable feedback to teachers that student test scores do not. Simply knowing students are (or are not) learning does not tell teachers much about how or what to improve. Observation, if done well, can provide specific and relevant feedback to teachers that can guide and focus their efforts to improve practices that have a direct impact on student achievement (Allen, Pianta, Gregory, Mikami, & Lun, 2011). Unfortunately, there are too few observational measures of teachers' performance that are grounded in developmental theory, hence relevant to teaching adolescents and applicable to the wide variation in classroom settings across content and ability. This study presents confirmatory evidence of the theoretical structure of an observational assessment of classroom interactions, the Classroom Assessment Scoring System (CLASS), across several samples of secondary school classrooms.

The need for measures of teacher performance with direct application in school settings is high. Unfortunately, most current systems for observing teachers are inadequate. As detailed in the 2009 report "The Widget Effect," 98% of teachers are given "satisfactory" ratings after being observed by principals. Most of the observational tools used to assess teachers are homegrown, with little to no empirical evidence that they can produce reliable metrics or are accurately measuring components of teaching actually associated with student learning (Pianta & Hamre, 2009). There are exceptions for

which empirical support has been obtained, most notably evaluation systems built from the Danielson protocol (Danielson, 2011), such as the Teacher Evaluation System in the Cincinnati Public Schools. When used systematically as a part of teacher evaluation or development systems, standardized and validated approaches to observation can improve teacher performance and student achievement for those teaching adolescents (Allen et al., 2011; Taylor & Tyler, 2011). Thus, harnessing the power of observational protocols to provide meaningful evidence from a teacher's classroom interactions is critical for the future of the educational system. An important aspect of rigorous evaluation of such protocols is evaluating the underlying factor structure inherent in observational measures, as these factor structures provide both a theoretical and descriptive basis for protocols, such as the CLASS.

Given the context of intense policy pressure on districts to adopt observational measures as a part of teacher evaluation systems and the developmental research on the environments in which adolescents best thrive, it is important that any assessment of teacher practices in secondary school classrooms be linked to an understanding of adolescent development. One observational measure used in several recent studies of middle and high school classrooms is the Classroom Assessment Scoring System–Secondary (CLASS-S; Pianta, Hamre, & Mintz, 2010). The CLASS-S is based on over a decade of research on the elements of teaching that contribute to students' cognitive and social development throughout school (Hamre & Pianta, 2010). The present article builds from recent work demonstrating associations between CLASS-S scores and students' achievement (Allen et al., 2011; Allen et al., 2013) and focuses more specifically on examining the factor structure of the CLASS-S across several large studies of middle and high school classrooms. We test the extent to which the previously theorized factor structure fit these diverse data sets and compare that fit with alternative structures. As discussed in greater detail throughout this article, this measurement work has important implications for research, policy, and practice as it provides empirical support to users in the interpretation of CLASS-S scores.

Theoretical Foundations of the Teaching Through Interactions Framework

Developmental theory and research provide strong support for the role of daily interactions of children and adolescents with adults and peers as the driving force behind learning and development (Bronfenbrenner & Morris, 1998). For example, a recent review of effective math programs in middle and high school provides support for an approach to understanding classroom quality that focuses on teacher-student interactions. Slavin, Groff, and Lake

(2009) reviewed evidence on 100 math interventions. Findings indicated that interventions focusing on daily interactions between teachers and students had stronger effects than did programs focusing solely on curricula and/or technology. This key role of interactions is somewhat in contrast to the current focus of middle and high schools in which content, standards, and curriculum are viewed as the most important elements of classrooms responsible for student learning (e.g., Confrey, 2006), suggesting a potentially dangerous lack of alignment.

In the present study, we discuss and test the factor structure of the proposed Teaching Through Interactions model (Hamre et al., 2013), using data collected with the CLASS-S observational measure. Consistent with many other descriptions from educational and psychological literature (e.g., Eccles & Roeser, 1999; Pressley et al., 2003; Roehrig et al., 2012), the Teaching Through Interactions model describes three broad domains of teacher-student interactions: Emotional Support, Classroom Organization, and Instructional Support. Each broad domain of teacher-student interactions consists of several specific dimensions of teacher-child interactions (see Table 1); for example, the domain of Emotional Support consists of four dimensions—Positive Climate, Negative Climate, Teacher Sensitivity, and Regard for Adolescent Perspectives. In effect, this model posits that effective teaching behaviors are organized into three broad domains, and these domains are applicable across grades and content area. A prior test of this three-domain conceptual model using data from over 4,000 pre-school to fifth-grade classrooms demonstrated the proposed model fit data well and provided superior fit to alternative one- or two-factor solutions (Hamre et al., 2013). However, this factor structure has yet to be confirmed in middle and high school classrooms where the organizing structure of teacher-student interactions may be different.

Below, we briefly review the theoretical background that led to the development of the Teaching Through Interactions framework and the observational tool on which it is based, the CLASS-S (Pianta et al., 2010). We then provide evidence supporting the validity of this approach to conceptualizing middle and high school classroom interactions by summarizing recent literature linking CLASS-S observations to student achievement.

Emotional Support

The Teaching Through Interactions framework describes four dimensions of teaching practice that are hypothesized to support students' social and emotional skills as well as their engagement in academic pursuits: Positive Climate, Negative Climate, Teacher Sensitivity, and Regard for Adolescent Perspectives (see Table 1 for operational definitions of these dimensions). One way the

Table 1. Description of Teaching Through Interactions at Secondary Level (Original CLASS-S Factors).

Domain	Dimension	Description
Emotional Support	Positive Climate	Reflects the emotional connection and relationships among teachers and students, and the warmth, respect, and enjoyment in verbal and non-verbal interactions.
	Negative Climate	Reflects the level and intensity of expressed negativity among teachers and students in the classroom.
	Teacher Sensitivity	Encompasses the teacher's awareness and responsiveness to the academic, social-emotional, and developmental needs of individual students and the entire class.
	Regard for Adolescent Perspectives	Focuses on the extent to which the teacher is able to meet and capitalize on the social and developmental needs of adolescents by providing opportunities for student autonomy and leadership. Also considered are the extent to which student ideas and opinions are valued and content is made useful and relevant to adolescents.
Classroom Organization	Behavior Management	Encompasses the teacher's use of clear behavioral expectations and effective methods to prevent and redirect misbehavior.
	Productivity	Considers how well the teacher manages time and routines so that instructional time is maximized.
	Instructional Learning Formats	Focuses on the ways in which the teacher maximizes students' interest and engagement in learning. This includes teachers' use of interesting and engaging lessons and materials, active facilitation, and clarity of learning objectives.
Instructional Support	Content Understanding	Refers to both depth of the lesson content and the approaches used to help students comprehend the framework, key ideas, and procedures in an academic discipline. At a high level, this refers to interactions among the teacher and students that lead to an integrated understanding of facts, skills, concepts, and principles.

(continued)

Table 1. (continued)

Domain	Dimension	Description
	Analysis and Inquiry	Assesses the degree to which the teacher facilitates students' use of higher-level thinking skills, such as analysis, problem solving, reasoning, and creation through the application of knowledge and skills. Opportunities for demonstrating meta-cognition, that is, thinking about thinking, are also included.
	Quality of Feedback	Assesses the degree to which feedback expands and extends learning and understanding and encourages student participation. (At the secondary level, significant feedback may be provided by peers)
	Instructional Dialogue	Captures the purposeful use of dialogue-structured, cumulative questioning and discussion that guide and prompt students—to facilitate students' understanding of content and language development. The extent to which these dialogues are distributed across all students in the class and across the class period is important to this rating.

Note. CLASS-S = Classroom Assessment Scoring System–Secondary.

CLASS-S is unique from versions of CLASS for earlier grade levels is that these emotional support dimensions are grounded in adolescent developmental literature. For example, a majority of what is coded in these dimensions are based on decades of research demonstrating that adolescents have particular developmental needs, including relational supports and connections, autonomy and competence, and relevance (Allen, Hauser, Bell, & O'Connor, 1994; Allen & Land, 1999; Ryan & Deci, 2000). Below, we review developmental research confirming the importance of each emotional support dimension.

Positive Climate

A strong student-teacher relationship is one key to positive academic performance, increased school motivation, and positive behavioral outcomes

(Rudasill, Gallagher, & White, 2010; Skinner, Furrer, Marchand, & Kindermann, 2008). It is well established that adolescents achieve more in challenging yet supportive environments to which they feel a positive connection (Eccles, 2004; Gentry, Gable, & Rizza, 2002). As a powerful example, one recent national study of almost 8,000 high school students demonstrated that schools in which students had more positive perceptions of the classroom relational climate had lower dropout rates (Barile et al., 2012). In contrast, adolescents who describe relationships with teachers as unsatisfying and unmotivating also report decreased levels of engagement in school (Roeser, Eccles, & Sameroff, 1998). Peer relationships and cooperation are also of vital importance to adolescents and need to be built into classroom interactions in meaningful and productive ways (Crosnoe, Cavanagh, & Elder, 2003).

Negative Climate

The absence of expressions of negative emotion is important for student learning. Attachment theorists posit that when adults provide emotional support in a predictable, consistent, and safe environment, children become more self-reliant and are able to take risks as they explore the world because they know that an adult will be there to help them if they need it (Bowlby, 1969).

Teacher Sensitivity

Sensitive teachers are attuned and responsive to the social, emotional, and academic needs of students in their classrooms, while maintaining a focus on the classroom as a whole. Students in classrooms with sensitive teachers are more engaged and self-reliant in the classroom, have lower levels of mother-reported internalizing problems (National Institute of Child Health and Human Development Early Child Care Research Network, 2003), and display greater gains in academic skills (Connor, Son, Hindman, & Morrison, 2005). Although the construct of teacher sensitivity has not often been used to study middle and high school teachers, it shares some elements of more frequently assessed elements of these environments such as a teachers' "withitness" (Kounin, 1970), although this construct has tended to apply exclusively to teachers' attunement to behavioral issues in the classroom.

Regard for Adolescent Perspectives

Middle and high school teachers' skills in supporting students' need for autonomy and decision making are critical to creating a classroom

environment that engages an adolescent (Hafen et al., 2012). For example, providing students with meaningful choices within the classroom has been shown to increase engagement (Allen et al., 1994). A mismatch between the adolescent's need for greater autonomy and the teacher's exercise of control has been shown to result in decreased student learning (Eccles, Wigfield, & Schiefele, 1998). This is particularly true of 6th- to 12th-grade classrooms due to the developmental needs for autonomy in guiding positive development. In other words, inattention to adolescent perspectives diminishes opportunity for student growth, while teacher-student interactions in which the student feels recognized as an individual with his or her own point of view enable an expansion of opportunity and motivation to learn and perform.

Classroom Organization

The Teaching Through Interactions framework describes three dimensions of teaching practice that are hypothesized to support students' abilities to regulate behavior and attention in the classroom such that they can get the most out of learning opportunities: Behavior Management, Productivity, and Instructional Learning Formats (see Table 1 for operational definitions of these dimensions).

Behavior Management

A multitude of studies indicate that classrooms with positive behavior management tend to have students who make greater academic progress (Tingstrom, Sterling-Turner, & Wilczynski, 2006). Middle and secondary classroom teachers who are effective classroom managers provide predictability allowing adolescents the chance to focus on learning; the establishment of a comfortable and orderly classroom where a teacher is proactively addressing student behavior promotes student engagement and leads to higher academic achievement (Woolfolk Hoy & Weinstein, 2006).

Productivity

Productive classrooms are those in which students are consistently exposed to learning opportunities and downtime is kept to a minimum. Classrooms and lessons tend to run smoothly when students are explicitly taught routines and procedures (Leinhardt, Weidman, & Hammond, 1987). With four to eight different classes daily, middle and secondary teachers must maximize learning time with lessons that get underway quickly, have a clear

beginning and end, and have very quick transitions between activities (Arlin, 1979).

Instructional Learning Formats

Student achievement is increased when teachers provide clear learning targets and specific feedback within an organized classroom with few behavioral disruptions (Brophy, 1986). Variety and novelty in modes of presentation and types of activities are also important teacher practices for student learning (Cotton, 2000). Yair (2000), for example, found that students in Grades 6 to 12 demonstrated low engagement during lectures but higher levels of engagement when participating in more active classroom activities (e.g., labs, groups). Thus, offering a more varied instruction that actively facilitates engagement is likely to lead to more learning in secondary school classrooms.

Instructional Support

The Teaching Through Interactions framework describes four instructional dimensions of teaching practice that are hypothesized to enhance students' cognition and learning: Content Understanding, Analysis and Inquiry, Quality of Feedback, and Instructional Dialogue (see Table 1 for operational definitions of these dimensions). Adolescents learn best in environments in which teachers hold high academic standards and expectations, but these must be accompanied by sufficient support for each individual to ensure that goals are attainable (Weinstein, Tomlinson-Clarke, & Curran, 2004). This is another area of the CLASS-S that is qualitatively different than earlier grades, as the extent to which the classroom environment pushes higher-order thinking and meta-cognition is of paramount importance in secondary school classrooms.

Content Understanding

Many elements of effective instruction promote students' understanding of material across content areas, and this requires connecting specific material to larger ideas. Teachers, therefore, need to organize instruction within a framework that builds toward the big ideas within an academic discipline but that is fully supported by a strong base of factual knowledge and skills (Bransford, Brown, & Cocking, 2000). Effective teachers teach subject matter in depth, providing many examples in which the same concept is at work and in which similarities and differences are explicitly addressed (Marzano, Pickering, & Pollock, 2001). Effective teachers present new material in small

steps, review relevant previous learning and prerequisite skills and knowledge (Bransford et al., 2000), and guide student performance through modeling, numerous examples, and opportunities for extensive practice, both supervised and independent (Swanson, Hoskyn, & Lee, 1999). Supervised practice of procedures and skills builds toward automaticity, thereby freeing up cognitive space for deeper understanding (Bransford et al., 2000). Content learning is strongest when teachers explicitly tie new information to students' background knowledge and real-world examples (Bransford et al., 2000) and when multiple perspectives are presented (Hooper & Rieber, 1995).

Analysis and Inquiry

Effective teachers tap into the natural problem-solving abilities and curiosity of students by providing them with opportunities to solve ill-defined problems and apply learning to novel contexts (Bransford et al., 2000). Instruction is organized to all students to engage in higher-order thinking (Marzano et al., 2001) and meta-cognitive processes (Bransford et al., 2000) that foster high levels of cognitive engagement and learning. Effective teachers intentionally and explicitly model thinking processes as a way of scaffolding students' development of meta-cognitive processes (Houtveen & van de Grift, 2007).

Quality of Feedback

High quality feedback also serves to enhance student learning either by bridging the gap between a student's current level and the target goal (e.g., scaffolding and hints) and/or by "pushing" the student to think or process information in greater depth (e.g., explaining or answering additional questions). Effective feedback is immediate, contingent, corrective and/or specific, and tied to natural settings (Marzano et al., 2001; Swanson et al., 1999). Such feedback serves to control frustration, increase interest and motivation and effort, and promote learning and higher-order thinking (Wood, Bruner, & Ross, 1976).

Instructional Dialogue

A final element of effective instructional support that is relevant across content areas concerns the dialogues that occur between teachers and students and among students. A large body of work demonstrates that students learn more when they are engaged in deep and meaningful conversations about content (Alexander, 2008). This literature highlights the importance

of building a shared dialogue, in contrast to the more typically seen classroom conversation patterns in which teachers ask a question, students respond, and teachers ask follow-up questions (Sinclair & Coulthard, 1975). Alexander (2008) also carefully distinguishes more casual instructional conversations from dialogues that are "characterized by purposeful questioning and chaining of ideas into covert lines of thinking and inquiry." These strategies may be particularly important to employ in the context of collaborative group work, with evidence that these collaborative groups promote better learning in classrooms in which teachers explicitly teach students about how to talk with one another (Mercer, Dawes, & Staarman, 2009).

Empirical Evidence Linking CLASS to Student Performance

There are now dozens of studies demonstrating that pre-school and elementary school-age children in classrooms with higher CLASS scores have better social and academic performance (see Downer, Sabol, & Hamre, 2010 for review). Students in classrooms with higher CLASS scores have been shown to have higher academic skills and achievement (Hamre, Hatfield, Pianta, & Jamil, 2014; La Paro, Pianta, & Stuhlman, 2004; Mashburn et al., 2008), less conflict with their teachers (Hamre & Pianta, 2005), and greater social skills (Burchinal, Vandergrift, Pianta, & Mashburn, 2010; Curby et al., 2009; Mashburn et al., 2008). There is less evidence regarding associations between CLASS-S and student outcomes; however, several recent studies provide important emerging evidence that the CLASS-S assesses components of middle and high school classrooms that are associated with students' academic and social development (Allen et al., 2011; Allen et al., 2013; Hafen et al., 2012; Kane et al., 2012).

The CLASS-S was one of the several observational measures used in the large Measures of Effective Teaching Study (Kane et al., 2012). Using CLASS-S data from math and English language arts teachers in Grades 4 through 8, the study demonstrated consistent associations between an overall composite of CLASS-S scores and student performance on state and national achievement tests. These links were stronger than those from teachers' years of experience and level of education. Additionally, students in classrooms with higher CLASS-S scores learned more and reported feeling more connected to school. Associations were stronger in math than in language arts classrooms.

Further evidence for the validity of CLASS-S was obtained in a much smaller sample of middle and high school teachers across multiple academic

content areas (Allen et al., 2013). This study used a state-specified achievement score as the outcome and all the dimensions of the CLASS-S observations as the predictor, controlling for each student's prior achievement. Effects ranged in size from very weak (.01) to moderate (.48) and were strongest for positive climate, teacher sensitivity, regard for adolescent perspectives, instructional learning formats, and analysis and inquiry. These effects were consistent for math, science, and English language arts (ELA) classrooms. The strongest support for the value of the teacher-student interactions, as measured by CLASS-S, comes from a coaching intervention (*MyTeachingPartner–Secondary*) study that used random assignment to demonstrate not only that teachers who were coached on the CLASS-S dimensions improved the effectiveness of their interactions with students but also that students in their classrooms learned more as a result (Allen et al., 2011; Gregory, Allen, Mikami, Hafen, & Pianta, 2014).

Factor Structure of the CLASS: The Current Study

These early validation studies also point to the need for greater clarity around the factor structure of the CLASS. Published studies by Allen and colleagues (2011) and Allen et al. (2013) and the Measures of Effective Teaching (MET) study (Kane et al., 2012) have each used different approaches to factoring CLASS-S data. In neither case did study authors use only the suggested and hypothesized factor structure discussed above. Allen and colleagues started by confirming the original three-factor structure but then created a single composite from a subset of five CLASS-S dimensions that were most strongly related to student achievement scores, while the MET study suggested that a single-factor solution was the most parsimonious. Although these approaches certainly have their utility, this presents an issue for researchers and practitioners. In an effort to provide researchers and practitioners a more general solution and empirical evidence to support decision making about how to best utilize CLASS-S data in researcher and policy settings, the present study tests a conceptually generated a priori structure across multiple data sets.

Recent studies of the classroom environment also suggest that there may be a global responsive teaching domain that is somewhat independent of the unique effects of emotional or instructional interaction quality. This approach takes advantage of advancements in quantitative methodology and the work around bi-factor models. The bi-factor model specifies both general and domain-specific factors that are uncorrelated and, thus, can be utilized as unique predictors in the same model (Chen, Hayes, Carver, Laurenceau, & Zhang, 2012). In recent work with the CLASS at the pre-K level, there is

evidence that a bi-factor model positing a general responsive teaching domain and more specific motivational, management, and cognitive facilitation domains fits the data somewhat better than a typical three-factor structure (Hamre et al., 2014) and is useful in examining associations with student learning gains. Thus, a bi-factor approach may provide an interesting alternative worthy of examination for CLASS-S.

In the current study, we will present results on various approaches for factor analysis and factor structure in an effort to understand the organization of teacher-student interactions in secondary classrooms. The current study systematically assesses the degree to which the hypothesized factor structure of the CLASS-S fits data from several studies of middle and high school classrooms. In a similar article conducted on data from over 4,000 pre-school to fifth-grade classrooms, the factor structure suggested by the Teaching Through Interactions framework fits the data better than alternative one- or two-factor models (Hamre et al., 2013). In this study, the fit of the hypothesized three-factor structure was examined relative to four alternative structures: a one-factor structure that loads all dimensions on a single factor, a two-factor model that combines emotional support and instructional support into one factor while leaving classroom organization as a separate factor, the original three-factor model, and a bi-factor approach that hypothesizes a global factor of responsive teaching with subfactors representing the three-factor approach.

The data set used for these analyses combines observations from four studies (Allen et al., 2011; Bell et al., 2012; Kane et al., 2012) of over 1,000 middle and high school classrooms across the country. These studies cover a broad array of classroom contexts, including schools in rural, suburban, and urban environments, as well as schools with diverse student populations. Thus, the observational data from these studies provide a robust test of applying a specific conceptual model of classroom settings to a broad spectrum of secondary classrooms in the United States. To test the applicability and generalizability of the Teaching Through Interactions' three-domain organization of teacher-student interactions, we used confirmatory factor analyses (CFA) to examine the extent to which the three-domain latent structure posited in the Teaching Through Interactions framework fits the natural variation in observed classroom processes in comparison with a single-factor, a two-factor, or a bi-factor solution. Consistent with the theoretical and empirical data cited above, we expected that the three-domain model, in which dimensions were organized under Emotional, Organizational, and Instructional interactions, would provide a better fit to the data than the alternative models, and that this data structure would remain similar across the different studies.

Table 2. Descriptive Data for Classroom, Teacher, and Double-Scored Reliability.

	Study			
	MTP-S	MET	TUCC	UTQ
Teachers				
N	67	875	82	458
Gender (in percent)				
Female	64	77	62	83
Male	36	23	38	17
Ethnicity (in percent)				
White (non-Hispanic)	84	58	20	57
African American	10	32	50	38
Hispanic	—	5	1	2
Asian American	—	—	26	2
Other	6	4	4	1
Mean years of experience	8.9 (9.3)	9.76 (8.7)	11.9 (9.0)	10.2 (7.3)
Rater reliability				
Weighted Kappa	.41	.29	.41	.14
Range	.31-.55	.17-.36	.22-.54	.08-.24
ICC	.42	.36	.43	.15
Range	.33-.51	.26-.49	.24-.58	.09-.25
Mean % within 1	81.3	77.3	79.9	81.9

Note. MTP = MyTeachingPartner; MET = Measures of Effective Teaching; TUCC = Toward an Understanding of Classroom Context; UTQ = Understanding Teaching Quality; ICC = intraclass correlations.

Method

Studies and Sample

The present study utilizes data from four large-scale, observational research projects conducted from 2007 to 2011 in 1,482 6th- through 11th-grade classrooms across the United States. Basic information for each study's sample and the double-scored reliability of the classroom observations are included in Table 2. Although the number of observations of each teacher varied across samples, a single average score for each teacher was used in analyses to avoid issues concerning independence of observations. Readers are referred to individual study citations for more complete information on the data collection procedures and sample.

Measures of Effective Teaching. The MET study (Kane et al., 2012) data were collected between 2010 and 2011. Videotapes of classroom interactions were

available from 875 teachers in Grades 6 to 8. The vast majority of teachers were videotaped four times throughout 1 academic year. Observers coded these tapes in 20-minute segments. Since the MET study was more representative demographically and comprised the largest number of teachers, it is used in analyses for both exploratory and confirmatory analyses.

MyTeachingPartner–Secondary (MTP-S). The MTP-S study (Allen et al., 2011) data were collected between 2007 and 2009. Videotapes of classroom interactions were available from 67 teachers in Grades 6 to 11. Teachers were videotaped for 40 minutes six times throughout 1 academic year. A total of 10 observers coded these tapes in 20-minute segments.

Toward an Understanding of Classroom Context (TUCC). The TUCC study (Bell et al., 2012) data were collected in 2009 and 2010. Videotapes of classroom interactions were available from 82 algebra teachers in 20 middle schools and 20 high schools. Teachers were videotaped four to five times throughout one academic year. A total of five observers coded these tapes in 15-minute segments.

Understanding Teaching Quality (UTQ). The UTQ study data were collected in 2009 and 2010. Videotapes of classroom interactions were available from 458 teachers in middle school mathematics and ELA classrooms. Teachers were videotaped four to five times throughout one academic year. A total of five observers coded these tapes in 15-minute segments.

Observational Measure

The CLASS-S (Pianta et al., 2010) was used to code all video observations. The CLASS-S version used in these studies has 10 dimensions. In each study, global ratings during each observation (ranging from 10 to 30 minutes) were made on a 7-point scale, ranging from "Low" to "High," for the following dimensions of teacher-student interactions: positive climate, teacher sensitivity, regard for adolescent perspectives, negative climate, behavior management, productivity, instructional learning formats, content understanding, analysis and inquiry, and quality of feedback. In the MET study only, there was one additional dimension (instructional dialogue). Observers watched a video of classroom interactions for a prescribed segment of time (anywhere from 10 to 30 minutes) while taking detailed field notes about specific teacher and student behaviors and interaction patterns. Observers then had 10 minutes to compare their notes with a CLASS-S manual and use the behaviorally anchored set of rating scales to record a final code for each dimension. Each session of coding is considered a cycle, which might consist of several

segments (e.g., 15-minute time periods). CLASS-S scores used in current analyses were aggregated across cycles, observers, and observation visits to form variables at the classroom level. This aggregation serves to increase the reliability of a teacher's CLASS-S score, as long as there is no evidence of meaningful variance across cycles.

Training and reliability. Across studies, all observers attended a workshop led by a CLASS-S certified trainer to attain initial reliability on the CLASS-S. The workshops consisted of guided practice with coding videotaped classroom footage. After the training workshops, observers had to pass a videotaped reliability test, involving either five or six cycles of 20 to 44 minutes. Criteria for passing required coders to score at least 80% match (within 1 scale point) with master codes on the global rating scales.

During data collection, all studies completed further reliability checks through independent dual review of videotapes. In the MTP-S study, the TUCC study, and the data used from the MET study, all observations were double-coded. In the UTQ study, 25% of the observations were double-coded. Reliability estimates from this double coding are provided in Table 2. In calculating weighted kappa, the estimates are derived at the segment level within dimensions and then averaged up to the dimension level. In calculating intraclass correlations (ICC), the estimates are averaged to the lesson level within dimensions.

Data Analysis Approach

Due to the relative size of the sample, data from the MET study were used for initial confirmatory and exploratory analyses. Data analysis was modeled after the approach of Hamre et al. (2013) and involved four steps to determine whether the original three-factor structure fit the data best or whether another structure provided a more accurate fit to the data: (a) a CFA on the large MET sample using the original three-factor structure; (b) an EFA on the MET sample to test whether another three-factor structure is more appropriate; (c) CFAs using the MET sample confirming whether a one-, two-, three-, or bi-factor latent structure provided a better fit to the MET data; and finally, (d) CFAs across the remaining data sets (MTP-S, TUCC, and UTQ) to confirm whether the findings from the MET sample replicate. It is of note that the dimension of instructional dialogue is only present in the MET sample.

Our evaluation criteria included a comparative fit index (CFI) above 0.90 and close to 0.95, root mean square error of approximation (RMSEA) under 0.10 and close to 0.06, and a standardized root mean square residual (SRMR) under 0.12 and close to 0.08 (Hu & Bentler, 1999; Schermelleh-Engel,

Moosbrugger, & Müller, 2003). Given that data are missing in some studies, we used full information maximum likelihood (FIML) as the method to estimate our structural parameters. Given that the amount of missingness is small for all of the studies (<5%) and that the results were consistent when these data were listwise deleted, it is unlikely that missing data affected the results. The benefits of maximum likelihood estimation with incomplete data have been extensively studied (e.g., Arbuckle, 1996). Multiple fit indices were then compared across the models, because the chi-square overall goodness-of-fit test is unfavorably affected by large sample size, model misspecification, or violation of distribution assumptions (Bollen, 1990).

Results

Comparing Three-Factor Models: The Theoretical Model Versus the Revised Model

The first analysis was intended to establish whether the theoretically derived three-factor model provided a good fit to the data. The largest sample (MET) was used in this analysis. In the theoretical model, there are three factors each with different dimensions: (a) Positive Climate (PC), Negative Climate (NC), Teachers' Sensitivity (TS), and Regard for Adolescent's Perspective (RAP) load on the Emotional Support factor; (b) Behavior Management (BM), Productivity (P), and Instructional Learning Format (ILF) load on the Classroom Organization factor; and (c) Content Understanding (CU), Analysis and Inquiry (AI), Quality of Feedback (QoF), and Instructional Dialogue (ID) load on the Instructional Support factor. We tested this structural hypothesis through a CFA framework. As can be seen in Table 3, this original three-factor model provided relatively poor fit to the data, $\chi^2(30) = 1044.34$, $p < .001$, CFI = 0.80, RMSEA = 0.24.

The second analysis was intended to explore whether an alternative factor structure provided a better fit to the data. The largest sample (MET) was again used in this analysis. In order to remove the redundancy of running this exploratory analysis on the same sample a follow-up confirmatory analysis would be conducted, we used a split-sample approach. Thus, we ran the EFA on only half of the sample. This EFA confirmed that a three-factor structure was best but with a slightly different factor structure than the original three-factor structure; thus, this alternative model was included in the confirmatory analyses. The new revised model was exactly the same as the original model except Negative Climate loaded most heavily on the Classroom Organization domain and Instructional Learning Formats loaded on the Instructional Support domain.

Table 3. Summary of Model Fit Indices as a Function of Study Samples and Alternative Confirmatory Factor Models.

	Single-factor model		Three-factor model original		Bi-factor model		Three-factor model revised	
Sample	CFI	RMSEA	CFI	RMSEA	CFI	RMSEA	CFI	RMSEA
MET	0.74	0.27	0.80	0.24	0.93	0.11	0.93	0.10
MTP-Secondary	0.74	0.27	0.84	0.16			0.92	0.09
TUCC	0.78	0.28	0.84	0.25			0.91	0.16
UTQ	0.71	0.22	0.74	0.21			0.88	0.15

Note. ES = Positive Climate, Teacher Sensitivity, and Regard for Adolescent Perspectives; CO = Negative Climate, Behavior Management, and Productivity; IS = Instructional Learning Formats, Content Understanding, Analysis and Inquiry, and Quality of Feedback. Residual correlations were included for the associations between Regard for Adolescent Perspectives and Analysis and Inquiry, as well as Regard for Adolescent Perspectives and Instructional Learning Formats. The bi-factor model is presented for the MET data set as a comparison and not for the other data sets as this approach is less optimal in smaller data sets. CFI = comparative fit index; RMSEA = root mean square error of approximation; MET = Measures of Effective Teaching; MTP = MyTeachingPartner; TUCC = Toward an Understanding of Classroom Context; UTQ = Understanding Teaching Quality.

In the third analysis, we tested this proposed three-factor structural hypothesis through a CFA framework. We ran the revised three-factor CFA on the half of the MET sample not included in the EFA. As can be seen in Table 3, this revised three-factor model provided a much better fit to the data, $\chi^2(30) = 643.31$, $p < .001$, CFI = 0.94, RMSEA = 0.10. The resulting factor loadings from this analysis are depicted in Table 4. All loadings were significant. Positive Climate, Teacher Sensitivity, and Regard for Student Perspectives loaded positively on Emotional Support. Behavior Management and Productivity loaded positively on Classroom Organization, while Negative Climate loaded negatively. Instructional Learning Formats, Content Understanding, Analysis and Inquiry, Quality of Feedback, and Instructional Dialogue all loaded positively on the Instructional Support factor.

As expected, each domain of teacher-student interactions was correlated with the other domains. The strongest association was between Emotional Support and Instructional Support ($\rho_{co,es} = .71$ [.02]). Emotional Support and Classroom Organization also were strongly correlated with each other ($\rho_{co,is} = .52$ [.02]), as were Classroom Organization and Instructional Support ($\rho_{es,is} = .59$ [.02]).

Reliability scores were computed for each of the three factors (see Geldhof, Preacher, & Zyphur, 2014). The formula for the reliability of the composite score is defined for the confirmatory factor model as

Table 4. Results of Confirmatory Revised Three-Factor Model.

	Emotional support		Classroom organization		Instructional support		Intercept		Residual variance	
Dimensions	Est	*SE*	Est	*SE*	Est	*SE*	Est	*SE*	Est	*SE*
Positive Climate	0.896	0.012					4.229	0.032	0.199	0.022
Teacher Sensitivity	0.882	0.014					4.130	0.025	0.222	0.023
Regard for Student Perspective	0.768	0.021					2.887	0.025	0.420	0.030
Negative Climate			−0.772	0.016			1.437	0.020	0.414	0.028
Behavior Management			0.949	0.009			5.896	0.034	0.066	0.016
Productivity			0.889	0.110			5.885	0.026	0.210	0.019
Instructional Learning Format					0.875	0.014	4.151	0.023	0.238	0.023
Content Understanding					0.799	0.018	3.729	0.021	0.367	0.028
Analysis and Inquiry					0.752	0.022	2.347	0.022	0.449	0.030
Quality of Feedback					0.868	0.015	3.430	0.026	0.245	0.024
Instructional Dialogue					0.822	0.019	3.156	0.025	0.287	0.025
Factor correlations										
Emotional Support	1.000	—								
Classroom Organization	0.599	0.032	1.000	—						
Instructional Support	0.798	0.012	0.648	0.030	1.000	—				

Note. Est = estimate; *SE* = standard error.

$$\rho_{xx} = \frac{\sum_{i=1}^{m}\lambda_i^2 + \sum_{i}^{m}\sum_{j}^{m}\lambda_i\lambda_j}{\sum_{i=1}^{m}\lambda_i^2 + \sum_{i}^{m}\sum_{j}^{m}\lambda_i\lambda_j + \sum_{i}^{m}\theta_{ii}}(i \neq j),$$

where m is the number of indicators and λ_I and θ_I are a factor loading and residual variance for the ith variable respectively. For the hypothesized three-factor model, reliability of Emotional Support was estimated at $\rho_{xx} = .72$. For Classroom Organization, $\rho_{xx} = .84$, and that of Instructional Support was computed as $\rho_{xx} = .91$. These internal consistency measures are all greater than an acceptable cutoff value of $\rho_{xx} = .70$.

Comparing One-, Two-, Three-, and Bi-Factor Models of Classroom Observational Data

Raw data for CLASS-S teacher-student interaction scores were directly fit to alternative confirmatory factor models. For the original and revised three-factor models, we structured the factors as described above. We then tested two alternate structures. In the single-factor model, all indicants loaded on a unitary factor. In the two-factor models, we tested all possible combinations

of combining two factors into one latent construct (e.g., Emotional Support and Instructional Support) and leaving the third factor isolated (e.g., Classroom Organization).

Table 3 shows fit indices for the one-factor model, the original three-factor model, the bi-factor model, and the revised three-factor model using the complete MET data. The most noticeable feature in the results is that the revised three-factor solution is the best fitting model among alternative models, as the combined evaluation across fit indices including CFI, RMSEA, and SRMR (not depicted in the table) indicated that the revised three-factor model provided the best fit. The single-factor model (RMSEA = 0.27), the three possible two-factor models (RMSEA = 0.24-0.33), and the original three-factor model (RMSEA = 0.24) all failed to be acceptable-fitting models. The bi-factor model (RMSEA = 0.11) was comparable with the revised three-factor model (RMSEA = 0.10), although it did not provide a better fit.

A final set of CFAs tested whether findings from the MET data would replicate in the other three samples (MTP-S, TUCC, and UTQ). The revised three-factor model was confirmed in each sample as the best fitting (see Table 3).

Follow-Up Analyses

A set of multiple-group CFAs were conducted to determine whether there were differences in the fit of the revised three-factor model based on known demographic characteristics in the MET sample. There were no significant differences in the fit of the revised three-factor model based on course subject (math, ELA), grade, or urbanicity of the school location (urban vs. rural).

Discussion

There is a clear need in educational practice for measures that combine developmental theory with classroom practices, as there is a need for clearer articulations and validation of higher-level theories of effective teaching (Douglas, 2009). This study provides a direct test of the Teaching Through Interactions theoretical model and finds some support for a structure that organizes teacher-student interactions in three factors: Emotional Support, Classroom Organization, and Instructional Support. Ultimately, empirical evidence supporting this theoretical structure has the potential to help researchers and practitioners decide how best to measure classroom interactions in secondary school classrooms in order to capture classroom environments that are most healthy for adolescents' social and academic development.

The findings from this study suggest that the three-factor structure providing the best fit to the data was only slightly different from the three-factor solution reported for elementary classrooms and yet was best described at the domain level in terms of Emotional Support (Positive Climate, Teacher Sensitivity, and Regard for Adolescent Perspectives), Classroom Organization (Behavior Management, Productivity, and Negative Climate), and Instructional Support (Instructional Learning Formats, Content Understanding, Analysis and Inquiry, Quality of Feedback, and Instructional Dialogue). When tested against other potential models for organizing classroom interactions, such as a one omnibus teacher quality factor or a simpler social and instructional supports model, the three-factor model had the best fit to data from over 1,000 secondary school classrooms. Importantly, this three-factor model fit observational data collected from a range of studies across a broad range of settings, including urban and rural classrooms, and across 6th- to 12th-grade classrooms.

As noted, this structure supported the model posited by the Teaching Through Interactions model with two exceptions: Negative Climate was best integrated in Classroom Organization instead of Emotional Support, and Instructional Learning Formats was organized within Instructional Support instead of Classroom Organization. The former makes logical sense, as when negative climate is high in a classroom, there tends to be disruptions to both the management of the classroom and the productivity of the classroom. The latter also is intuitive as the variety of formats presented in instruction are directly related to the extent to which instruction is varied and emphasizing higher-order thinking (Bransford et al., 2000; Marzano et al., 2001). Although this revised organization is slightly different from the original structure that has been reliably confirmed in pre-K and early elementary samples, it is likely a reflection of the typically more rigidly structured environment in middle and high school classrooms. For instance, scores for Behavior Management and Productivity for CLASS-S tend to be negatively skewed with a mean near 6 on a scale from 1 to 7, while scores for those same dimensions in pre-K classrooms tend to be more uniformly distributed with a mean closer to 5.5.

This study suggests that the proposed factor structure of the CLASS-S is relevant across grade levels, diverse contexts, and across content areas, although specific confirmatory tests for grade level were not conducted. A theory and measure for effective teaching that can be applied so widely, particularly when taken together with validity studies documenting effects of these types of interactions, may be important for scaling and application.

The implications of this study for informing researchers and practitioners are notable. For researchers, as a broad framework based on classroom interactions

allows us to better consolidate literature on effective teaching—rather than fragmenting that literature by what may be somewhat artificial distinctions such as grade level or content area (Hamre et al., 2013). However, researchers should not feel confined in their use of the CLASS-S dimensions to the factor structure presented here. Although this study provides evidence that the revised three-domain structure best reflects interactions across the many classrooms studied, there are alternative approaches that may be valuable for varying research contexts and questions. For instance, there may be good reasons to focus efforts on particular dimensions in work that is aimed toward professional development (e.g., Allen et al., 2011). When the goal is improving classroom interactions in order to improve student achievement, a targeted focus on those dimensions most strongly related to achievement is likely beneficial. Also, there is recent evidence that statistical approaches that separate broad frameworks into factor structures that are uncorrelated (e.g., bi-factor models) may allow for greater predictive validity (Hamre et al., 2014). This approach is particularly useful for researchers who not only attempt to isolate the unique effects of the global domain of responsive teaching but are also interested in the underlying effects of more specific interactions (e.g., academic-cognitive facilitation). Additionally, the statistics from this study indicate the bi-factor model fits the data about as well as the revised three-factor structure. However, the use of the bi-factor model for practitioners may be quite difficult given its empirical rather than theoretical basis. Further, the bi-factor model is more susceptible to bias in smaller data sets and often presents convergence issues. Thus, the revised three-factor model tested is likely to remain the predominantly utilized structure for CLASS-S.

For practitioners, the CLASS-S observational protocol has already been translated into an implementation guide describing the dimensions in detail along with suggestions of how to implement those practices more often in their instruction. Emphasizing the importance of interactions to teachers is a powerful tool. Teachers often consider improving the day-to-day interactions they have with their students as a means for improving student learning, but have few tools or guides for describing or improving these behaviors, even though there is ample evidence to suggest their importance (Slavin et al., 2009). Further, the practical use of the CLASS-S framework has already been demonstrated with its inclusion in effective professional development (Allen et al., 2011; Mikami, Gregory, Allen, Pianta, & Lun, 2011). For practitioners, in the context of understanding classroom interactions, it allows for a consistent description of quality interactions that can form the basis of a shared community of practice within schools and across disciplines. It also has the benefit of face validity as teachers can readily understand what it means to implement high quality emotional, organization, and instructional interactions in the classroom.

Another area of research in which the Teaching Through Interactions framework and the three domains of classroom interactions presented in this study may enhance our knowledge is in capturing how teachers develop specific skills over time. In a study that followed teachers from their last year of teacher education into their first 2 years of teaching, evidence showed some distinct patterns of development across teachers (Malmberg, Hagger, Burn, Mutton, & Colls, 2011). Specially, results indicated that, on average, teachers increased their classroom organization and management skills over these early years of teaching, with the strongest improvements seen among teachers who started off relatively lower in this domain. A different pattern was found for emotional support, where initial increases were followed by declines, leading to an inverted U-shape curve over time. This nuanced view of teacher development is only possible when researchers consider multiple domains of teaching quality rather than overall changes over time based on student test scores or teacher report (Hamre et al., 2013).

The goal of any measurement work in the field of education should ultimately lie in promoting positive changes in education practice. Indeed, there is general consensus within the educational community that the professional development of teachers is of paramount importance (Caspary, 2002). However, professional development typically occurs in the absence of a direct link to actual teaching behavior in classrooms, particularly for already-trained and certified teachers (Caspary, 2002). Systematic classroom observation systems provide a standardized approach to measuring and noting teachers' strengths and weaknesses and evaluating whether professional development activities are actually helping improve the classroom interactions responsible for learning. Along with continuing to develop professional development programs, future research might productively be geared toward exploration of the ways in which these findings could translate into day-to-day practices of teacher assessment and training. For example, further research might consider whether it was possible to conduct CLASS-S ratings via live observation by regular school personnel, as part of ongoing teacher development efforts.

Limitations

The study is not without limitations. First, the inter-rater reliabilities for the rating scales used in this study, while considered acceptable, were not as high as are found in less comprehensive or more discrete measures of teacher behavior. Recent work has suggested that there may be some utility in breaking apart an observational measure into sources of variance using Generalizability Theory (e.g., Casabianca et al., 2013). In this work, error

associated with rater, day, and setting can be isolated and incorporated to improve reliability of measurement. This is an area of need across many observational measures although it was beyond the scope of the present study. Second, this study tested a theoretical model of teaching using an observational measure that was developed from a specific theory. It will be important for future work to examine the extent to which data from measures not only based on different theoretical frameworks but also focusing on teacher-student interactions (e.g., Danielson, 2011) may also fit this broad conceptual model for the way teacher-student interactions are organized. Finally, it is quite likely that that measurement of teacher-student interactions are influenced by characteristics of individual students or the group more generally, and so interpretations of these results as indicative of only the teacher are unwarranted. Future research could help determine whether the demographic composition of students in a classroom (e.g., % low income, % ethnic minority) significantly alters the proposed structure of classroom interactions.

Conclusion

The Teaching Through Interactions framework offers only one view into the quality of teacher-student interactions in classrooms, interactions that are among the most important aspects of teachers' job. There is also initial evidence that this is a robust model for understanding classroom interactions consistent across culture (Malmberg & Hagger, 2009) and grade level (Allen et al., 2011; Mashburn et al., 2008). This study confirms that teacher-student interactions, as observed by the CLASS-S can be reasonably organized into domains of emotional, organizational, and instructional behavior across a range of samples. In doing so, this framework provides a vehicle that may help link basic research to teacher evaluation and professional development.

Declaration of Conflicting Interests

The author(s) declared the potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Bridget K. Hamre and Robert C. Pianta are authors of the CLASS measure and own a company that conducts CLASS trainings.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study and its write-up were supported by grants from the Bill and Melinda Gates Foundation, the William T. Grant Foundation, and the Institute for Education Science (R305A100367).

References

Alexander, R. (2008). Culture, dialogue, and learning: Notes on an emerging pedagogy. In N. Mercer & S. Hodgkinson (Eds.), *Exploring talk in school: A celebration of the work of Douglas Barnes* (pp. 91-114). London, England: SAGE.

Allen, J. P., Gregory, A., Mikami, A. Y., Lun, J., Hamre, B. K., & Pianta, R. C. (2013). Observations of effective teaching in secondary school classrooms: Predicting student achievement with the CLASS-S. *School Psychology Review*, *42*, 76-98.

Allen, J. P., Hauser, S. T., Bell, K. L., & O'Connor, T. G. (1994). Longitudinal assessment of autonomy and relatedness in adolescent-family interactions as predictors of adolescent ego development and self-esteem. *Child Development*, *65*, 179-194.

Allen, J. P., & Land, D. (1999). Attachment in adolescence. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment theory and research* (pp. 319-331). New York, NY: Guilford.

Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science*, *333*, 1034-1037.

Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 243-277). Mahwah, NJ: Lawrence Erlbaum.

Arlin, M. (1979). Teacher transitions can disrupt time flow in classrooms. *American Education Research Journal*, *16*, 42-56.

Barile, J. P., Donohue, D. K., Anthony, E. R., Baker, A. M., Weaver, S. R., & Henrich, C. C. (2012). Teacher-student relationship climate and school outcomes: Implications for educational policy initiatives. *Journal of Youth and Adolescence*, *41*, 256-267.

Bell, C. A., Gitomer, D. H., McCaffrey, D. F., Hamre, B. K., Pianta, R. C., & Qi, Y. (2012). An argument approach to observation protocol validity. *Educational Assessment*, *17*, 62-87.

Bollen, K. A. (1990). *Structural equations with latent variables*. New York, NY: John Wiley.

Bowlby, J. (1969). *Attachment and loss: Attachment* (Vol. 1). New York, NY: Basic Books.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

Bronfenbrenner, U., & Morris, P. A. (1998). The ecology of developmental processes. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology—Vol. 1: Theoretical models of human development* (5th ed., pp. 993-1029). New York, NY: John Wiley.

Brophy, J. (1986). Teacher influences on student achievement. *American Psychologist*, *41*, 1069-1077.

Burchinal, M., Vandergrift, N., Pianta, R. C., & Mashburn, A. (2010). Threshold analysis of association between child care quality and child outcomes for low-income

children in pre-kindergarten programs. *Early Childhood Research Quarterly*, *25*, 166-176.

Casabianca, J. M., McCaffrey, D. F., Gitomer, D. H., Bell, C. A., Hamre, B. K., & Pianta, R. C. (2013). Effect of observation mode on measures of secondary mathematics teaching. *Educational and Psychological Measurement*, *73*, 757-783.

Caspary, K. (2002). California's pioneering training and retention initiatives for early childhood educators. Berkeley, CA: PACE.

Chen, F. F., Hayes, A., Carver, C. S., Laurenceau, J.-P., & Zhang, Z. (2012). Modeling general and specific variance in multifaceted constructs: A comparison of the bi-factor model to other approaches. *Journal of Personality*, *80*, 219-251.

Confrey, J. (2006). The evolution of design studies as methodology. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 135-152). New York, NY: Cambridge University Press.

Connor, C. M., Son, S., Hindman, A. H., & Morrison, F. J. (2005). Teacher qualifications, classroom practices, and family characteristics: Complex effects on first graders' language and early reading. *Journal of School Psychology*, *43*, 343-375.

Cotton, K. (2000). *The schooling practices that matter most*. Portland, OR: Northwest Regional Educational Laboratory and the Association for Supervision and Curriculum Development.

Crosnoe, R., Cavanagh, S., & Elder, F. H. (2003). Adolescent friendships as academic resources: The intersection of friendship, race, and school disadvantage. *Sociological Perspectives*, *46*, 331-352.

Curby, T. W., LoCasale-Crouch, J., Konold, T. R., Pianta, R., Howes, C., Burchinal, M., . . .Barbarin, O. (2009). The relations of observed pre-k classroom quality profiles to children's academic achievement and social competence. *Early Education and Development*, *20*, 346-372.

Danielson, C. (2011). Evaluations that help teachers learn. *Educational Leadership*, *4*, 35-40.

Douglas, K. (2009). Sharpening our focus in measuring classroom instruction. *Educational Researcher*, *38*, 518-521.

Downer, J. T., Sabol, T., & Hamre, B. K. (2010). Teacher-child interactions in the classroom: Toward a theory of within- and cross-domain links to children's developmental outcomes. *Early Education and Development*, *21*, 699-723.

Eccles, J. S. (2004). Schools, academic motivation, and stage-environment fit. In R. M. Lerner & L. D. Steinberg (Eds.), *Handbook of adolescent psychology* (2nd ed., pp. 125-153). New York, NY: John Wiley.

Eccles, J. S., & Roeser, R. W. (1999). School and community influences on human development. In M. H. Boorstein & M. E. Lamb (Eds.), *Developmental psychology: An advanced textbook* (4th ed., pp. 503-554). Hillsdale, NJ: Lawrence Erlbaum.

Eccles, J. S., Wigfield, A., & Schiefele, U. (1998). Motivation to succeed. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology* (5th ed., Vol. III, pp. 1017-1095). New York, NY: John Wiley.

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods*, *19*, 72-91.

Gentry, M., Gable, R. K., & Rizza, M. G. (2002). Students' perceptions of classroom activities: Are there grade-level and gender differences? *Journal of Educational Psychology*, *94*, 539-544.

Gregory, A., Allen, J. P., Mikami, A. Y., Hafen, C. A., & Pianta, R. C. (2014). Effects of a professional development program on behavioral engagement of students in middle and high school. *Psychology in the Schools*, *51*, 143-163.

Hafen, C. A., Allen, J. P., Mikami, A. Y., Gregory, A., Hamre, B., & Pianta, R. C. (2012). The pivotal role of adolescent autonomy in secondary school classrooms. *Journal of Youth and Adolescence*, *41*, 245-255.

Hamre, B. K., Hatfield, B., Pianta, R. C., & Jamil, F. (2014). Evidence for general and domain-specific elements of teacher-child interactions: Associations with preschool children's development. *Child Development*, *85*, 1257-1274.

Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first grade classroom make a difference for children at risk of school failure? *Child Development*, *76*, 949-967.

Hamre, B. K., & Pianta, R. C. (2010). Classroom environments and developmental processes: Conceptualization and measurement. In J. Meece & J. Eccles (Eds.), *Handbook of research on schools, schooling, and human development* (pp. 25-41). New York, NY: Routledge.

Hamre, B. K., Pianta, R. C., Downer, J. T., DeCoster, J., Mashburn, A. J., & Hamagami, A. (2013). Teaching through interactions: Testing a developmental framework of teacher effectiveness in over 4,000 classrooms. *The Elementary School Journal*, *113*, 461-487.

Hooper, S., & Rieber, L. (1995). Teaching with technology. In A. C. Ornstein (Ed.), *Theory into practice* (pp. 155-170). Boston, MA: Allyn & Bacon.

Houtveen, T., & van de Grift, W. (2007). Reading instruction for struggling learners. *Journal of Education for Students Placed at Risk*, *12*, 405-425.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55.

Kane, T., Staiger, D., McCaffrey, D., Cantrell, S., Archer, J., Buhayar, S., & Parker, D. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains* (Technical report). Seattle, WA: Bill & Melinda Gates Foundation, Measures of Effective Teaching Project.

Kounin, J. S. (1970). *Discipline and group management in classrooms*. New York, NY: Holt, Rinehart, and Winston.

La Paro, L. M., Pianta, R. C., & Stuhlman, M. (2004). Classroom Assessment Scoring System (CLASS): Findings from the pre-K year. *Elementary School Journal*, *105*, 409-426.

Leinhardt, G., Weidman, C., & Hammond, K. M. (1987). Introduction and integration of classroom routines by expert teachers. *Curriculum Inquiry*, *17*, 135-176.

Malmberg, L.-E., & Hagger, H. (2009). Changes in student teachers' agency beliefs during a teacher education year, and relationships with observed classroom quality, and day-to-day experiences. *British Journal of Educational Psychology*, *79*, 677-694.

Malmberg, L.-E., Hagger, H., Burn, K., Mutton, T., & Colls, H. (2011). Observed classroom quality during teacher education and two years of professional practice. *Journal of Educational Psychology*, *102*, 916-932.

Marzano, R., Pickering, D., & Pollock, J. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement*. Alexandria, VA: Association for Supervision and Curriculum Development.

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O., Bryant, D., . . .Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, *79*, 732-749.

Mercer, N., Dawes, L., & Staarman, J. K. (2009). Dialogic teaching in the primary science classroom. *Language and Education*, *23*, 353-369.

Mikami, A. Y., Gregory, A., Allen, J. P., Pianta, R. C., & Lun, J. (2011). Effects of a teacher professional development intervention on peer relationships in secondary classrooms. *School Psychology Review*, *40*, 367-385.

NICHD Early Child Care Research Network. (2003). Social functioning in first grade: Prediction from home, child care, and concurrent school experience. *Child Development*, *74*, 1639-1662.

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, *38*, 109-119.

Pianta, R. C., Hamre, B. K., & Mintz, S. (2010). *Classroom Assessment Scoring System–Secondary (CLASS-S)*. Charlottesville: University of Virginia.

Pressley, M., Roehrig, A., Raphael, L., Dolezal, S., Bohn, C., Mohan, L., . . .Hogan, K. (2003). Teaching processes in elementary and secondary education. In W. Reynolds & G. Miller (Eds.), *Handbook of psychology: Educational psychology* (Vol. 7., pp. 153-176). Hoboken, NJ: John Wiley.

Roehrig, A. D., Turner, J. E., Arrastia, M. C., Christesen, E., McElhaney, S., & Jakiel, L. M. (2012). Effective teachers and teaching: Characteristics and practices related to positive student outcomes. In K. R. Harris, S. Graham, & T. Urdan (Eds.), *APA Educational Psychology Handbook: Individual Differences and Cultural and Contextual Factors* (Vol. 2, pp. 501-527). Washington, DC: American Psychological Association.

Roeser, R. W., Eccles, J. S., & Sameroff, A. J. (1998). Academic and emotional functioning in early adolescence: Longitudinal relations, patterns, and prediction by experience in middle school. *Development and Psychopathology*, *10*, 321-352.

Rudasill, K. M., Gallagher, K., & White, J. M. (2010). Temperamental attention and activity, classroom emotional support, and academic achievement in third grade. *Journal of School Psychology*, *48*, 113-134.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*, 68-78.
Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research*, *8*, 23-74.
Sinclair, J. M., & Coulthard, R. M. (1975). *Towards an analysis of discourse: The English used by teachers and pupils*. London, England: Oxford University Press.
Skinner, E. A., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of Educational Psychology*, *100*, 765-781.
Slavin, R., Groff, C., & Lake, C. (2009). Effective programs in middle and high school mathematics: A best-evidence synthesis. *Review of Educational Research*, *79*, 839-911.
Swanson, H., Hoskyn, M., & Lee, C. (1999). *Interventions for students with learning disabilities*. New York, NY: Guilford.
Taylor, E. S., & Tyler, J. H. (2011). *The effect of evaluation on performance: Evidence from student achievement data of mid-career teachers* (NBER Working Paper No. 16877). Cambridge, MA: National Bureau of Economic Research.
Tingstrom, D. H., Sterling-Turner, H. E., & Wilczynski, S. M. (2006). The good behavior game: 1969-2002. *Behavioral Modification*, *30*, 225-253.
Weinstein, C. S., Tomlinson-Clarke, S., & Curran, M. (2004). Toward a conception of culturally responsive classroom management. *Journal of Teacher Education*, *55*, 25-38.
Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, *17*, 89-100.
Woolfolk Hoy, A. W., & Weinstein, C. S. (2006). Student and teacher perspectives on classroom management. In C. S. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice, and contemporary issues* (pp. 181-219). Mahwah, NJ: Lawrence Erlbaum.
Yair, G. (2000). Educational battlefields in America: The tug-of-war over students' engagement with instruction. *Sociology of Education*, *73*, 247-269.

Author Biographies

Christopher A. Hafen is currently a research scientist at the University of Virginia's Center for Advanced Study of Teaching and Learning (CASTL). He received his PhD in developmental psychology from Florida Atlantic University and is the manager of an educational intervention project (www.myteachingpartner.com). His research interests include understanding the function and development of close relationships in adolescence.

Bridget K. Hamre is currently an associate research professor and associate director of the University of Virginia's CASTL. She received her PhD from the University of Virginia. Her research interests include student-teacher relationships and classroom

processes that promote positive academic and social development for young children. She has recently worked with leaders in several states and the Office of Head Start to implement CLASS as a tool to enhance teacher-child interactions through accountability and professional development systems.

Joseph P. Allen is a Hugh P. Kelly professor of psychology at the University of Virginia. He has done extensive research in the area of adolescent social development, with an emphasis on using observational techniques to identify the critical facets of adolescents' relationships with others in their social world (both adults and peers). He has also conducted randomized trials of interventions targeting school failure, school dropout, and other social difficulties (e.g., teen pregnancy) among adolescents.

Courtney A. Bell is a senior research scientist at Educational Testing Service (ETS) in the Understanding Teaching Quality Center. A former high school science teacher and teacher educator, her research focuses on the intersections of policy and practice with a particular focus on the measurement and assessment of teaching, teacher education, and teaching policy. Her most recent mixed methods studies investigate how raters make sense of teaching through observational protocols and how administrators learn to use observation protocols.

Drew H. Gitomer is the Rose and Nicholas DeMarzo Chair in education at Rutgers University. He received his PhD from the University of Pittsburgh. His research centers on the assessment and evaluation of teaching and related policy issues in teaching and teacher education. He is involved in work carrying out validity studies of a variety of measures, including classroom observation protocols, classroom assignment protocols, and new measures of teacher knowledge.

Robert C. Pianta is currently dean and Novartis professor of education for the Curry School of Education at the University of Virginia. He received his PhD from the University of Minnesota. His work focuses on policy and practice that enhance children's outcomes, school readiness, and later achievement. He has served on the steering committee and was a study investigator for the NICHD Early Child Care Research Network.